(6)

REGENERATIVE SIMULATION FOR ESTIMATING EXTREME VALUES

by

Donald L. Iglehart and Mark L. Stone

TECHNICAL REPORT NO. 64

July 1982

Prepared under Contract N00014-76-C-0578 (NR 042-343)*

for the

Office of Naval Research

DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

*This research was also partially supported under
National Science Foundation Grant MCS79-09139.

REGENERATIVE SIMULATION FOR ESTIMATING EXTREME VALUES

by

Donald L. Iglehart

and

Mark L. Stone
Stanford University

1. Introduction

Let $\underset{\sim}{X} = \{X_t : t \geq 0\}$ be a regenerative process which we wish to simulate. Under mild regularity conditions the distribution of X_t converges to the distribution of some limiting random variable (or vector) X. This type of convergence is known as weak convergence and written $X_t \Rightarrow X$, as $t \uparrow \infty$. Simulators speak of X as the "steady-state" configuration of the system and are often interested in estimating the constant $r = E\{f(X)\}$, where f is a given real-valued function defined on the state-space of the process $\underset{\sim}{X}$. The regenerative method of estimation provides a means of constructing point and interval estimates for r; see IGLEHART (1977) for an expository summary of this method.

The problem we consider in this paper does not involve estimation of r, but rather the estimation of extreme values of the regenerative process $\underset{\sim}{X}$. Suppose, for the sake of discussion, we are simulating a stable GI/G/1 queue in order to estimate the maximum waiting time among the first $n+1$ customers; call this random variable W_n^*. As n grows, so will W_n^*. However, W_n^* does not converge to a finite limit, but rather diverges to $+\infty$. We will be interested in estimating the distribution function of W_n^* for finite, but large n. By the same token we might wish to estimate the distribution of the maximum queue length during the interval $[0,t]$. While this problem of estimating extreme values would seem to be of great practical importance to simulators, we know of no papers in the simulation literature which offer any guidance on the subject. This paper will attempt to partially fill the gap.

We begin in Section 2 by summarizing a series of probabilistic results in extreme value theory which will provide the theoretical basis for the methods we propose. Section 3 discusses several methods for estimating extreme values for the general regenerative simulation. In Sections 4 and 5 we treat the special cases of the GI/G/1 queue and birth-death processes respectively. Theoretical results are available for these two classes of regenerative processes that are useful in assessing the accuracy of the estimation methods proposed. Section 6 contains the numerical results for simulations of the M/M/1 queue carried out to illustrate the estimation methods proposed.

2. Probabilistic Background

Let $\{F_n : n \geq 1\}$ be a sequence of distribution functions (d.f.'s) on the real line, $\mathbb{R} = (-\infty, +\infty)$. This sequence converges weakly to a d.f. G if $\lim_{n \to \infty} F_n(x) = G(x)$ for all $x \in \mathbb{R}$ which are continuity points of G. We write $F_n \Rightarrow G$ to denote this type of convergence. If X_n (resp. X) is a random variable (r.v.) with d.f. F_n (resp. G), we also write $X_n \Rightarrow X$ to denote this weak convergence. Sometimes it is convenient to write $X_n \Rightarrow G$ to denote the same thing. The material presented in this section can be found for the most part in deHAAN (1970), currently the best comprehensive treatment of the subject.

Now let $\{X_n : n \geq 1\}$ be a sequence of independent, identically distributed (i.i.d.) r.v.'s and denote the maximum of the first n r.v.'s by $M_n = \max\{X_j : 1 \leq j \leq n\}$. If each of the X_j's has d.f. F, then M_n will have d.f. F^n. We shall say that F belongs to the domain of attraction of the nondegenerate d.f. G, and write $F \in \mathcal{D}(G)$, if we can choose two sequences of constants $\{a_n : n \geq 1\}$ and $\{b_n : n \geq 1\}$ with $a_n > 0$ such that

$$F^n(a_n x + b_n) \to G(x) \tag{2.1}$$

as $n \to \infty$ for all $x \in \mathbb{R}$ for which G is continuous. Equivalently, $F \in \mathcal{D}(G)$ if $(M_n - b_n)/a_n \Rightarrow G$ as $n \to \infty$. Thus for large n we would approximate $P\{M_n \leq x\}$ by $G((x-b_n)/a_n)$. If a r.v. X has d.f. $F \in \mathcal{D}(G)$, we also write $X \in \mathcal{D}(G)$.

A famous result in extreme value theory states that the only d.f.'s G which can arise in (2.1) are of one of the following three types:

$$\Phi_\alpha(x) = \begin{cases} 0\,, & x \leq 0 \\ \exp(-x^{-\alpha})\,, & x > 0 \end{cases} \tag{2.2}$$

$$
(2.3) \qquad \Psi_\alpha(x) = \begin{cases} \exp\{-(-x)^\alpha\}, & x < 0 \\ 1, & x \geq 0 \end{cases}
$$

$$
(2.4) \qquad \Lambda(x) = \exp(-e^{-x}), \qquad x \in \mathbb{R},
$$

where in (2.2) and (2.3) α is a positive constant. Recall that two d.f.'s G_1 and G_2 are said to be of the same type if there exist two constants a and b, $a > 0$, such that $G_1(x) = G_2(ax + b)$ for all $x \in \mathbb{R}$. Thus aside from translations and scaling by a positive constant the three d.f.'s given in (2.2) - (2.4) are the only ones that can appear in (2.1). This result on the three types of limit d.f.'s is usually attributed to GNEDENKO (1943), however it was first formulated in this way by FISHER and TIPPETT (1928).

The next logical result to seek is necessary and sufficient conditions for $F \in \mathcal{D}(G)$, where G of necessity is one of the three d.f.'s given in (2.2) - (2.4). Furthermore, if $F \in \mathcal{D}(G)$ we need a method for selecting the two sequences $\{a_n : n \geq 1\}$ and $\{b_n : n \geq 1\}$. To this end we first define the right endpoint, $x_0 \leq +\infty$, of the d.f. F as

$$
x_0 = \sup\{x : F(x) < 1\} .
$$

A d.f. $F \in \mathcal{D}(\Phi_\alpha)$ if and only if for all $x > 0$

$$
(2.5) \qquad \lim_{t \to \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-\alpha} .
$$

If $F \in \mathcal{D}(\Phi_\alpha)$, then we can take

$$a_n = \inf\{x : 1 - F(x) \leq 1/n\} \tag{2.6}$$

and $b_n = 0$. A d.f. $F \in \mathcal{D}(\Psi_\alpha)$ if and only if $x_0 < \infty$ and for all $x > 0$

$$\lim_{t \to \infty} \frac{1 - F[x_0 - (tx)^{-1}]}{1 - F(x_0 - t^{-1})} = x^{-\alpha} \ . \tag{2.7}$$

If $F \in \mathcal{D}(\Psi_\alpha)$, then we can take $b_n = x_0$ and

$$a_n = x_0 - \inf\{x : 1 - F(x) \leq 1/n\} \ .$$

The final case, $F \in \mathcal{D}(\Lambda)$, is the most important one for our simulation applications. A d.f. $F \in \mathcal{D}(\Lambda)$ if and only if

$$\lim_{t \uparrow x_0} \frac{1 - F(t + xf(t))}{1 - F(t)} = e^{-x} , \qquad \text{for all } x \in \mathbb{R} , \tag{2.8}$$

where for $t < x_0$

$$f(t) = \frac{\int_t^{x_0} (1 - F(s))ds}{1 - F(t)} \ .$$

If $F \in \mathcal{D}(\Lambda)$, then we can take

$$b_n = \inf\{x : 1 - F(x) \leq 1/n\}$$

and

$$a_n = \frac{\int_{b_n}^{x_0} [1 - F(t)]dt}{1 - F(b_n)} \quad .$$

Alternative expressions are available for a_n and b_n. We can use a'_n and b'_n provided $a_n/a'_n \to 1$ and $(b_n - b'_n)/a_n \to 0$. Let $Q_n(p)$ denote the p-quantile of the d.f. F^n: for $0 < p < 1$,

$$Q_n(p) = \inf\{x : F^n(x) \geq p\} \quad .$$

Then if $F \in \mathcal{D}(\Lambda)$, we can alternatively select

$$(2.9) \qquad b_n = Q_n(e^{-1})$$

and

$$(2.10) \qquad a_n = Q_n(e^{-e^{-1}}) - Q_n(e^{-1}) \quad .$$

Furthermore, if $F \in \mathcal{D}(\Lambda)$ and $x_0 = +\infty$, $M_n/b_n \Rightarrow 1$ as $n \to \infty$. Many of the classical d.f.'s such as the exponential, gamma, normal, lognormal, and logistic belong to $\mathcal{D}(\Lambda)$.

Suppose F has $x_0 = +\infty$ and possesses an exponential tail:

$$(2.11) \qquad 1 - F(x) \sim b \exp(-ax) \; , \qquad \text{as} \quad x \to \infty \; ,$$

where a and b are two positive constants. Then it is easy to check that (2.8) holds and $F \in \mathcal{D}(\Lambda)$. Using the expressions (2.9) and (2.10) it can be shown that b_n and a_n can be selected as follows:

(2.11a) $$b_n = a^{-1} \ln(nb) ,$$

and

(2.11b) $$a_n = a^{-1} .$$

An interesting (and practical) situation arises if F is a discrete d.f. as, for example, the geometric d.f. $F(x) = 1 - \exp(-[x])$, $x \geq 0$, where $[x]$ is the integer part of x. In this case neither (2.5), nor (2.8) hold, and since $x_0 = +\infty$, F does not belong to the domain of attraction of any of the three types (2.2) - (2.4). However, a result has been salvaged by ANDERSON (1970). Let $\mathcal{A}$ be the class of all d.f.'s whose support consists of all sufficiently large positive integers. Then for $F \in \mathcal{A}$,

(2.12) $$\limsup_{n \to \infty} F^n(\alpha^{-1}x + b_n) \leq \exp(-e^{-x})$$

and

(2.13) $$\liminf_{n \to \infty} F^n(\alpha^{-1}x + b_n) \geq \exp(-e^{-(x-\alpha)})$$

for some $\alpha > 0$, all x, and some sequence $\{b_n : n \geq 1\}$ if and only if

(2.14) $$\lim_{n \to \infty} \frac{1 - F(n)}{1 - F(n+1)} = e^{\alpha} .$$

When this condition holds, the constants b_n can be selected as follows. For $F \in \mathcal{A}$ and each positive integer n let $h(n) = -\log(1-F(n))$ and define h_c to be the extension of h obtained by linear interpolation for $x \geq 1$. Then define for $x \geq 1$

$$F_c(x) = 1 - \exp(-h_c(x)) \quad .$$

Clearly F_c is a continuous d.f. and for sufficiently large x is strictly increasing since $F \in \mathcal{A}$. For $x \leq 1$ the F_c can be defined arbitrarily just so long as it is a d.f. In terms of F_c we can define b_n for large n as the unique root of

$$1 - F_c(b_n) = 1/n \quad .$$

If $F \in \mathcal{A}$ and $1 - F(n) \sim b \exp(-an)$ as $n \to \infty$ $(a,b > 0)$, then for $b_n = a^{-1} \ln(nb)$ it can easily be shown using the method followed by HEYDE (1970) that for integer ℓ

$$\lim_{n \to \infty} \left[P\{M_n - [b_n] \leq \ell\} - \exp(-e^{-a(\ell-d_n)}) \right] = 0 , \tag{2.15}$$

where $d_n = b_n - [b_n]$. Thus for n large we would (ignoring a possible continuity correction) use the approximation

$$P\{M_n \leq \ell + [b_n]\} \cong \exp(-e^{-a(\ell - d_n)}) \quad ,$$

or

$$P\{M_n \leq \ell\} \cong \exp(-e^{-a(\ell - b_n)}) \quad . \tag{2.16}$$

Suppose now that we also have defined on the probability triple $(\Omega, \mathfrak{F}, P)$ that supports the i.i.d. sequence $\{X_n : n \geq 1\}$ a renewal process $\{\ell(t) : t \geq 0\}$ with mean time between renewals m $(0 < m < \infty)$. Then the weak law of large numbers for renewal processes states that $\ell(t)/t \Rightarrow m^{-1}$ as $n \to \infty$. Next set

$$M_t' = \max\{X_j : 1 \leq j \leq \ell(t)\} \quad .$$

The following useful result for this situation was obtained by BERMAN (1962). If $(M_n - b_n)/a_n \Rightarrow G$, one of the three extreme value d.f.'s (2.2) - (2.4), then as $t \to \infty$

$$(M_t' - b_{[t]})/a_{[t]} \Rightarrow G^{1/m} \quad . \tag{2.17}$$

This result provides a useful tool for extreme values of regenerative processes. To be explicit suppose $\underset{\sim}{X} = \{X_t : t \geq 0\}$ is a regenerative process defined on $(\Omega, \mathfrak{F}, P)$ and T_j, $j \geq 1$, is the time of the jth regeneration point of $\underset{\sim}{X}$ with $T_0 = 0$. Then the renewal process $\{\ell(t) : t \geq 0\}$ which counts the number of regeneration points in $(0,t]$ is defined by

$$\ell(t) = \max\{j : T_j \leq t\}$$

with $\ell(0) = 0$. For $j \geq 1$, let

$$M_j^+ = \sup\{X_t : T_{j-1} \leq t < T_j\} .$$

Since $\underset{\sim}{X}$ is regenerative, the sequence of maxima, $\{M_j^+ : j \geq 1\}$, will be i.i.d. Then if $L_t = \sup\{X_s : 0 \leq s \leq t\}$, clearly

$$\text{(2.18)} \qquad \max\{M_j^+ : 1 \leq j \leq \ell(t)\} \leq L_t \leq \max\{M_j^+ : 1 \leq j \leq \ell(t) + 1\} .$$

Combining the inequalities of (2.18) with the limit theorem of (2.17) enables us to show that

$$\text{(2.19)} \qquad (L_t - b_{[t]})/a_{[t]} \Rightarrow G^{1/m} ,$$

where $m = E\{T_1\}$, provided $M_1^+ \in \mathcal{D}(G)$. Of course if $M_1^+ \in \mathcal{Q}$, then the weaker results of Anderson or Heyde are all that can be expected.

We conclude this section by summarizing the problems confronting us for a regenerative processes with continuous state space. If $M_1^+ \in \mathcal{D}(G)$, then we can use (2.19) to obtain the asymptotic (for large t) approximation

$$\text{(2.20)} \qquad P\{L_t \leq x\} \cong G^{1/m}\left(\frac{x - b_{[t]}}{a_{[t]}}\right) .$$

Alternatively, we can also show that

$$(M_{\ell(t)} - b_{\ell(t)})/a_{\ell(t)} \Rightarrow G$$

which when combined with (2.18) yields the asymptotic (for large t) approximation

$$P\{L_t \leq x\} \underset{=}{\sim} G\left(\frac{x - b_{\ell(t)}}{a_{\ell(t)}}\right) . \tag{2.21}$$

This expression does not require an estimate for m as is the case with (2.20).

If the simulation is run for n cycles, then (2.20) or (2.21) should be replaced by

$$P\left\{ \max_{1 \leq j \leq n} M_j^+ \leq x \right\} \cong G\left(\frac{x-b_n}{a_n}\right) . \tag{2.22}$$

For (2.20), (2.21), or (2.22) to be useful, we must estimate a_n, and b_n. To use (2.20) we must also estimate m. The expected cycle length, m, can of course be estimated by the sample mean of the cycle lengths observed. Several methods for estimating a_n and b_n will be discussed in Section 3. Finally, we must assess whether $M_1^+ \in \mathcal{D}(G)$ for one of the three d.f.'s (G's) given in (2.2) - (2.4). For many simulations in which extreme values are being estimated, the limit d.f.'s G will be either Λ or Φ_α, since the maxima arising are unbounded $(x_0 = +\infty)$. Our experience with specific examples indicates that if the regenerative process is stable (converges to a non-degenerate limit), then $G = \Lambda$; while if the process is "null-recurrent" $(m = E\{T_1\} = +\infty)$, then $G = \Phi_\alpha$. However, in this case $G^{1/m}(x) = 1$ for all $x > 0$ which indicates that a different normalization must be used to obtain a nondegenerate limit. In any case, we note that if $X \in \mathcal{D}(\Phi_\alpha)$ with constants $a_n' > 0$ and $b_n' = 0$, then $\ln X \in \mathcal{D}(\Lambda)$ with constants $a_n = \alpha^{-1}$ and $b_n = \ln a_n'$. We note in passing that the extreme value behavior of some functions of a regenerative process can be handled in the same way. If the state space of the regenerative process is discrete, then we shall only consider the situation in which the d.f. of $M_1^+ \in \mathcal{Q}$ and $P\{M_1^+ > n\} \sim b \exp(-an)$ as $n \to \infty$ for some $a,b > 0$. In this case we can approximate the d.f. of L_t or $\max\{M_j^+ : 1 \le j \le n\}$ by using (2.12) and (2.13) or (2.15) and (2.16).

3. Statistical Estimation Problem

We now present some procedures which can be used to estimate the d.f. of extreme values occurring in a regenerative stochastic process belonging to a (known) domain of attraction. (The methods will be illustrated for the case of maximum values, but analogous procedures could be used for minimum values.) The basic idea underlying all the procedures is to simulate a process in order to form an empirical d.f. of say the maximum of the process over a given period of time or number of cycles, and to regress this against the functional form of the appropriate extreme value distribution. The methods differ as to how the empirical d.f. is formed and how the regression is done. The applicability of the methods depends on the process lying in the domain of attraction of a nondegenerate d.f. For a process taking on positive values, but having right endpoint $x_0 < \infty$, convergence to a nondegenerate extreme value distribution would not hold. This would be the case for instance for the queue length process in a system with finite waiting room.

3.1. The Continuous State Space Case

Suppose we are interested in the maximum of a continuous state space regenerative process over n cycles. Choose an integer $k > n$, (the choice of k will be discussed later) and simulate k cycles of the process, yielding individual cycle maxima $y_{1,k} > y_{2,k} > \dots > y_{k,k}$. A sample of k cycles contains $\binom{k}{n}$ subsets of n cycles. We then use the simulation results to find the maximum of the process over each subset of n cycles. The only $y_{i,k}$'s which can be maxima for some subset of n cycles are $y_{1,k},\dots,y_{k-n+1,k}$. As explained below, we are only interested in $y_{1,k},\dots,y_{i_0,k}$ where the predetermined number i_0 is usually considerably less than $k-n+1$. Thus in performing the simulation, we need only keep track of $y_{1,k},\dots,y_{i_0,k}$, and the continuous state space assumption can be relaxed to the distinctness of $y_{1,k},\dots,y_{i_0,k}$. (For instance the customer waiting time process of a queueing system may have an atom at 0; however this usually need not concern us.) Now observe that $y_{i,k}$ is the maximum value in $\binom{k-i}{n-1}$ of the $\binom{k}{n}$ subsets of n cycles, $i = 1,\dots,k-n+1$. Thus $y_{1,k}$ is the maximum value in a fraction n/k of the $\binom{k}{n}$ cycles, $y_{2,k}$ is the maximum value in a fraction $n(k-n)/(k(k-1))$ of the $\binom{k}{n}$ cycles, etc. Let $E_n(x)$ (suppressing k) be the empirical d.f. of the maximum of the process over n cycles. Then

$$E_n(y_{1,k}) = 1,$$

$$E_n(y_{2,k}) = 1 - n/k = \frac{k-n}{k},$$

$$E_n(y_{3,k}) = \frac{k-n}{k} - \frac{n(k-n)}{k(k-1)} = \frac{k-n}{k} \cdot \frac{k-n-1}{k-1},$$

and generally $E_n(y_{i,k}) = E_n(y_{i-1,k}) \cdot \dfrac{k-n-i+2}{k-i+2}$ for $i = 2,\dots,i_0$. Thus $E_n(y_{i,k})$ can be obtained by a simple recursive computation.

It is desirable to have the values $E_n(y_{i,k})$ diminish slowly enough from 1 to 0 so as to provide good upper quantile information for the regression procedure. Thus we want a rather high k/n ratio (say, at least 10). But n should be large enough so that the extreme value distribution provides a good approximation to the maximum of the process over n cycles. So $E_n(y_{i,k})$ is very small for many of the latter values of the $k-n+1$ possible i's. Using too many $y_{i,k}$'s in the regression would tend to drown out the effect of the first few, and could be a source of noise when $E_n(y_{i,k})$ is very close to 0. Thus we only consider $y_{i,k}$ for $i \leq i_0$, where $i_0 = \max\{i : E_n(y_{i,k}) \geq \epsilon\}$ for ϵ such as .001.

Consider the nonlinear regression problem of selecting a_n and b_n so as to obtain the least-squares fit of $\Lambda((x_i - b_n)/a_n)$ to $E_n(x_i)$, when Λ is the appropriate domain of attraction, for data points $x_i = y_{2,k},\ldots, y_{i_0,k}$. For short we shall denote this problem by

$$E_n(x_i) \approx \exp(-\exp(-(x_i - b_n)/a_n)).$$

When Φ_α is the appropriate domain of attraction, the problem becomes

$$E_n(x_i) \approx \exp(-(x_i/a_n')^{-\alpha}).$$

(Note that the data form a dependent sample.) The point $y_{1,k}$ is not used in the regression since $E_n(y_{1,k}) = 1$, which effectively makes $y_{1,k}$ the right endpoint x_0, even though Λ has a right endpoint of $+\infty$. This manifests itself in the regression by forcing b_n to ∞ or a_n to 0. Refer to the above procedure (for a fixed k,n, and $i_0(\epsilon)$) as the <u>basic</u>

nonlinear regression procedure. In the case of domain Λ, this procedure provides estimates $\hat{a}_n$, $\hat{b}_n$ of the norming constants for Λ. An alternative procedure, referred to as the basic linear regression procedure, is to (take a double $-\ln$ transformation and) perform the linear regression $-\ln(-\ln(E_n(x_i))) \approx c_n x_i + d_n$, with x_i as above, to obtain estimates $\hat{c}_n$ and $\hat{d}_n$. This fit can be used directly, or equivalently note that we can form estimates $\hat{a}_n = \hat{c}_n^{-1}$, $\hat{b}_n = -\hat{c}_n^{-1}\hat{d}_n$ of the norming constants for Λ.

Recall that if $X \in \mathcal{D}(\Phi_\alpha)$ with parameters $a_n' > 0$ (and $b_n' = 0$) then $\ln X \in \mathcal{D}(\Lambda)$ with norming constants $a_n = \alpha^{-1}$ and $b_n = \ln a_n'$. Thus performing the nonlinear regression procedure on X, based on $X \in \mathcal{D}(\Phi_\alpha)$, would lead to $E_n(x_i) \approx \exp(-(x_i/a_n')^{-\alpha})$ in order to estimate a_n', α. Performing the nonlinear regression procedure on $\ln X$, based on $\ln X \in \mathcal{D}(\Lambda)$, would lead to $E_n(\ln x_i) \approx \exp(-\exp(-(\ln x_i - b_n)/a_n))$ in order to estimate a_n, b_n. On noting that $E_n(\ln x_i) = E_n(x_i)$ by strict monotonicity of $\ln$, it is apparent that the results are the same whether we perform the (nonlinear) procedure for X based on $X \in \mathcal{D}(\Phi_\alpha)$ or on $\ln X$ based on $\ln X \in \mathcal{D}(\Lambda)$. The same is true for the linear procedure since $E_n(x_i) \approx \exp(-(x_i/a_n')^{-\alpha})$ is transformed into

$$-\ln(-\ln(E_n(x_i))) \approx \alpha \ln x_i - \alpha \ln a_n'.$$

If one were unsure as to which (or any) domain of attraction is appropriate, regressions corresponding to all the candidate extreme value distributions could be performed and the adequacy of the fits evaluated. For instance, we could perform one regression using $y_{i,k}$'s and another using $\ln y_{i,k}$'s.

Consider the specialization to $\mathcal{D}(\Lambda)$ and assume the exponential tail assumption (2.11) holds. We can use either the basic linear or the nonlinear regression procedure to obtain estimates $\hat{a}_n$, $\hat{b}_n$ if n is large enough for the approximation (2.22) to be good. Then without further simulation or regression, we can use (2.11a) and (2.11b) to obtain estimates $\tilde{a}_{n'} = \hat{a}_n$, $\tilde{b}_{n'} = \hat{a}_n \ln(n'/n) + \hat{b}_n$. The linear or nonlinear regression procedure generally increases in accuracy with k/n, but the relation (2.22) deteriorates as n decreases. Thus for a fixed amount of simulation k, n should be chosen as small as the adequacy of (2.22) allows.

The relations (2.11a) and (2.11b) suggest a possible variation to either the basic linear or nonlinear regression procedure: simulate k cycles and substitute $a_n = a^{-1}$, $b_n = a^{-1}\ln(nb)$, in the right-hand side of the regression. Now form the empirical d.f. as in the basic procedure for a fixed $n = n_j$. Do this for several values of n_j, pool the sample, and perform one regression to estimate a^{-1}, b. Assuming (2.11a) and (2.11b) hold adequately, this variation extracts more information from a given amount of simulation. This comes at the cost of increased complexity, and has not been attempted.

3.2 Modifications for the Discrete State Space Case

Now assume the regenerative process of interest has a discrete state space and $M_1^+ \in \mathcal{A}$. If the basic (continuous state space) procedures were used, there could now be ties among the $y_{i,k}$'s. So we simulate k cycles of the process and let $y_{1,k} > y_{2,k} > \ldots > y_{\ell,k}$ be the distinct individual cycle maxima, occurring respectively in $N_{1,k},\ldots,N_{\ell,k}$ different cycles. The number of $N_{i,k}$'s could be quite small even for k large. Letting $N_{0,k} = 0$ and $m_{i,k} = \sum_{\ell=0}^{i-1} N_{\ell,k}$, $y_{i,k}$ is the maximum value in

$$\sum_{j=1}^{N_{i,k}} \binom{k-m_{i,k}-j}{n-1}$$

of the $\binom{k}{n}$ subsets of n cycles, for $i \le \ell$. Thus $E_n(y_{1,k}) = 1$, and if for instance $N_{1,k} = 2$, then $E_n(y_{2,k}) = \frac{k-n}{k} \cdot \frac{k-n-1}{k-1}$, etc. So $E_n(y_{i,k})$ can still be obtained by a simple recursion. In practice, we would only consider $y_{i,k}$ for $i \le i_0$, where $i_0 = \max\{i : E_n(y_{i,k}) \ge \epsilon\}$ for some small ϵ. Now i_0 is a function of the sample path since the number of duplications is not known a priori. Call the procedure where we throw out $y_{1,k}$ and do a regression as in the basic (continuous state space case), the basic full delete procedure (either linear or nonlinear). The basic partial delete procedure may not completely throw out $y_{1,k}$. Instead, we consider a modified empirical d.f. as follows: if $N_{1,k} \ge 2$ and say for example $N_{1,k} = 3$, then take $E_n(y_{1,k}) = \frac{k-n}{k} \cdot \frac{k-n-1}{k-1}$ and $E_n(y_{2,k}) = \frac{k-n}{k} \cdot \frac{k-n-1}{k-1} \cdot \frac{k-n-2}{k-2}$, etc. We only alter $E_n(y_{1,k})$, leaving $E_n(y_{i,k})$, $i > 1$ as before. If $N_{1,k} = 1$, the basic full and partial delete procedures are equivalent; otherwise they differ only in the use of an additional regression data point by the partial delete procedure. Most of the variations mentioned in the continuous state space case are also available in the discrete case.

After having obtained estimates $\hat{a}_n$, $\hat{b}_n$, we would approximate $P\{\max_{1 \le j \le n} M_j^+ \le x\}$ by $\exp(-\exp(-\hat{a}_n^{-1}(x + \delta - \hat{b}_n)))$ where δ represents a continuity correction such as $\delta = 0$ or $1/2$. Estimating the d.f. of $\max_{1 \le j \le n} M_j^+$ is tougher in the discrete than the continuous case for the following reasons: (i) a continuity correction (δ) is needed, (ii) true convergence to the extreme value distribution is not attained, suggesting that the approximation on which the regression equation is based may not be as good as in the continuous case, and (iii) due to duplications associated with the $y_{i,k}$'s, the number of data points for the regression is greatly reduced relative to the continuous case, for a given k,n, and cutoff ϵ.

The basic procedures are illustrated in Section 6 for the M/M/1 queue. Before leaving this section we point out some other relevant references. The reliability theory literature contains many references on the problem of testing whether observations come from an exponential or extreme value d.f. and of estimating the associated parameters. Two useful places to find such papers are EPSTEIN (1960) and MANN, SCHAFER, and SINGPURWALLA (1974), Chapter 5. PICKANDS (1975) has developed a method for determining which G d.f. is appropriate for a given set of observations. His method uses a random and increasing number of the $y_{j,n}$'s as n increases. The method is expensive computationally and emphasizes an aspect of the extreme value problem which is not of great concern for simulation. Finally, WEISSMAN (1978) contains another method for estimating the constants a_n and b_n.

4. The GI/G/1 Queue

The GI/G/1 queue and the birth-death processes treated in Section 5 are among the very few regenerative processes for which we know the values of a_n and b_n. For this reason these processes are excellent candidates for testing the effectiveness of the estimation procedures proposed in Section 3.

In the GI/G/1 queue we assume customer 0 arrives at $t_0 = 0$, finds a free server, and experiences a service time v_0. Customer n arrives at time t_n and experiences a service time v_n. Customers are served in their order of arrival and the server is never idle if customers are waiting. Let the interarrival times $t_n - t_{n-1} = u_n$, $n \geq 1$. We assume the two sequences $\{v_n : n \geq 0\}$ and $\{u_n : n \geq 1\}$ each consist of i.i.d. r.v.'s and are themselves independent. Let $E\{u_n\} = \lambda^{-1}$ and $E\{v_n\} = \mu^{-1}$, where $0 < \lambda, \mu < \infty$. The traffic intensity $\rho = \lambda/\mu$ is assumed to be less than one. We exclude the deterministic system in which both the v_n's and u_n's are degenerate. Let the waiting time of the nth customer be W_n, the workload (or virtual waiting time) at time t be V_t, and the number of customers in the system at time t be Q_t. Also set $W_n^* = \max\{W_j : 0 \leq j \leq n\}$, $V_t^* = \sup\{V_s : 0 \leq s \leq t\}$, and $Q_t^* = \sup\{Q_s : 0 \leq s \leq t\}$. Let $X_n = v_{n-1} - u_n$, $n \geq 1$, and set $S_n = X_1 + \cdots + X_n$ for $n \geq 1$ and $S_0 = 0$. If n_j denotes the number of customers served in the jth busy period, then n_1 is related to the partial sum process $\{S_n : n \geq 0\}$ since

$$n_1 = \inf\{n > 0 : S_n \leq 0\} .$$

When $\rho < 1$, we have $m = E\{n_1\} < \infty$. Also $-S_{n_1}$ is the length of the first idle period. We assume that X_1 has an aperiodic d.f. (support is <u>not</u> concentrated on a set of points of the form $0, \pm h, \pm 2h, \pm 3h, \ldots$), that there exists a positive number κ such that $E\{\exp(\kappa X_1)\} = 1$, and $0 < \mu_\kappa = E\{X_1 \exp(\kappa X_1)\} < \infty$. These assumptions will normally be satisfied if the d.f. of v_0 has an exponentially decaying tail; e.g., when v_0 has a gamma distribution. Under these conditions we know (see IGLEHART (1972)) that

$$(4.1) \qquad (W_n^* - \kappa^{-1} \ln c_1 n)/\kappa^{-1} \Rightarrow \Lambda^{1/m}(x) .$$

and

$$(4.2) \qquad (V_t^* - \kappa^{-1} \ln c_2 t)/\kappa^{-1} \Rightarrow \Lambda^{\lambda/m}(x) ,$$

where

$$c_1 = \frac{[1 - E\{e^{\kappa S_{n_1}}\}]^2}{\kappa\mu_\kappa m}$$

and

$$c_2 = E\{e^{\kappa v_0}\} c_1 .$$

Thus to use (4.1) and (4.2) for estimating the d.f.'s of W_n^* and V_t^* we need only estimate m and $E\{e^{\kappa S_{n_1}}\}$, assuming that κ, μ_κ, and $E\{e^{\kappa v_0}\}$ can be calculated numerically. In the special case of M/G/1 queues no estimation is required, since $m = (1-\rho)^{-1}$ and $E\{e^{\kappa S_{n_1}}\} = \lambda/(\lambda+\kappa)$. If the simulation is carried out for a fixed number of cycles, then counterparts of (4.1) and (4.2) hold with the exponents of Λ removed.

The queue-length process $\{Q_t : t \geq 0\}$ is discrete-valued and the associated d.f. of $M_1^+ \in \mathcal{Q}$. Hence a limit theorem comparable to (4.1) or (4.2) does not exist. Instead we must seek results like (2.12) and (2.13) or (2.15) and (2.16). Unfortunately, these results are only known for the M/G/1 and GI/M/1 queues; see COHEN (1969), Theorems 7.2 and 7.5. Let $M_j^+ = \sup\{Q_t : T_{j-1} \leq t < T_j\}$. Then for an M/G/1 queue the counterpart of (2.16) is

(4.3) $$P\{\max_{1 \le j \le n} M_j^+ \le \ell\} \cong \exp(-e^{-a(\ell-b_n)}) \quad ,$$

where

$$a = \ln((\lambda+\kappa)/\lambda)$$

and

$$b_n = a^{-1} \ln(c_2 n) \quad .$$

On the other hand, for GI/M/1 queues (4.3) holds with $a = \ln((\mu-\kappa)/\mu)$ and the same value for b_n. Tables 1 and 2 contain the values of m, κ, μ_κ, c_1, and c_2 for the M/M/1 and $M/E_2/1$ queues as a function of the traffic intensity ρ.

(4.4) EXAMPLE. M/M/1 queue. For this queue we are able to calculate exactly the distribution of M_1^+ associated with the waiting time process, the virtual waiting time process, and the queue length process. The tail of the distribution of M_1^+ for the waiting time process is given in COHEN (1969), p. 606, or can be calculated directly from results in IGLEHART (1972), p. 630. The result is

$$P\{M_1^+ > x\} = \frac{\rho(1-\rho)e^{-kx}}{1-\rho^2 e^{-kx}} \sim \rho(1-\rho)e^{-kx} \quad \text{as} \quad x \uparrow \infty \; .$$

The corresponding result for the virtual waiting time process is given in

TABLE 1

Parameter Values for M/M/1 Queue with $\mu = 10$

ρ	m	κ	μ_κ	c_1	c_2
.1	1.11	9	.900	.09	.9
.2	1.25	8	.400	.16	.8
.3	1.43	7	.233	.21	.7
.4	1.67	6	.150	.24	.6
.5	2.00	5	.100	.25	.5
.6	2.50	4	.067	.24	.4
.7	3.33	3	.043	.21	.3
.8	5.00	2	.025	.16	.2
.9	10.00	1	.011	.09	.1
.95	20.00	0.5	.005	.0475	.05
.99	100.00	0.1	.001	.0099	.01

TABLE 2

Parameter Values for $M/E_2/1$ Queue with $\mu = 10$

ρ	m	κ	μ_κ	c_1	c_2
.1	1.11	15.00	.3375	.1562	2.5
.2	1.25	12.60	.2016	.2346	1.7119
.3	1.43	10.61	.1395	.2874	1.3038
.4	1.67	8.83	.1012	.3179	1.0201
.5	2.00	7.19	.0741	.3263	0.7957
.6	2.50	5.64	.0534	.3118	0.6050
.7	3.33	4.16	.0367	.2733	0.4357
.8	5.00	2.73	.0227	.2094	0.2809
.9	10.00	1.35	.0106	.1188	0.1366
.95	20.00	0.67	.0051	.0630	0.0674
.99	100.00	0.13	.0010	.0132	0.0134

[4], p. 606, or can again be calculated from [10], p. 632. The result is

$$P\{M_1^+ > x\} = \frac{(1-\rho)e^{-kx}}{1-\rho e^{-kx}} \sim (1-\rho)e^{-kx} \quad \text{as} \quad x \uparrow \infty .$$

The corresponding result for the queue length process is given in Example (5.4). ◀

5. Birth-Death Processes

A second class of regenerative processes for which theoretical results are available is birth-death processes in discrete or continuous time. Let $\{X_n : n \geq 0\}$ be a discrete time Markov chain with state space $E = \{0, 1, 2, \ldots\}$ and transition probabilities given by

$$(5.1) \qquad p_{ij} = \begin{cases} q_i, & j = i-1 \\ p_i, & j = i+1 \\ 0\ , & \text{other } j, \end{cases}$$

where $q_0 = 0$, $p_0 = 1$ and the other q_i's and p_i's are positive. This chain will automatically be both irreducible and periodic. Furthermore, recall that it will be recurrent if and only if

$$\sum_{j=1}^{\infty} (\pi_j\ p_j)^{-1} = \infty$$

where $\pi_0 = 1$ and $\pi_j = (p_0 \cdots p_{j-1})/(q_1 \cdots q_j)$. We assume the chain is recurrent. It will be positive recurrent if and only if

$$\sum_{j=0}^{\infty} \pi_j < \infty \ .$$

Next define

$$\tau_1(k) = \inf\{n > 0 : X_n = k\}, \qquad k \in E \ ,$$

the first entrance time to state k. Let $P_i\{\cdot\} = P\{\cdot | X_0 = i\}$, the conditional probability of an event, given $X_0 = i$. Then our concern here will be in the probability, given $X_0 = i$, of the Markov chain entering state n before it enters state 0. Let this probability be denoted by

$$r_i(n) = P_i\{\tau_1(n) < \tau_1(0)\} \ , \qquad i \in \{1, 2, \ldots, n-1\} \ .$$

Fortunately, this probability has been calculated and in particular

$$r_0(n) = r_1(n) = \Big(1 + \sum_{i=1}^{n-1} (\pi_i p_i)^{-1}\Big)^{-1} \ ;$$

see CHUNG (1960), p. 68. Note that $\lim_{n \to \infty} r_0(n) = 0$ when the chain is recurrent, in keeping with our intuition. Define

$$M_1^+ = \sup\{X_n : 0 \le n \le \tau_1(0) - 1\} \ .$$

Then

$$(5.2) \qquad P_0\{M_1^+ > n\} = r_0(n+1) = \left(\sum_{i=0}^{n} (\pi_i p_i)^{-1}\right)^{-1} , \qquad \text{as } n \to \infty \ .$$

Suppose now that we have a birth-death process $\{X_t : t \geq 0\}$: a continuous time Markov chain with state space $E = \{0, 1, \ldots\}$ and embedded jump chain whose probabilities are given by (5.1). As above, define the first entrance time to state k and the maximum in the first cycle by

$$\tau_1(k) = \inf\{s > 0 : X_{s-} \neq k, X_s = k\}$$

and

$$M_1^+ = \sup\{X_t :\ 0 \leq t < \tau_1(0)\} .$$

Because of the path structure of the birth-death process, it is easy to see from (5.2) that

$$P_0\{M_1^+ > n\} = \left[1 + \lambda_0 \sum_{i=1}^{n} (\pi_i \lambda_i)^{-1}\right]^{-1} , \tag{5.3}$$

where λ_i [resp. μ_i] are the birth [resp. death] parameters and $\pi_0 = 1$, $\pi_i = (\lambda_0 \lambda_1 \cdots \lambda_{i-1}/\mu_1 \mu_2 \cdots \mu_i)$. The same argument can of course be used to show that (5.3) also holds for semi-Markov processes with embedded jump chain whose probabilities are given by (5.1).

(5.4) EXAMPLE. M/M/s queue. The queue-length process, $\{Q_t : t \geq 0\}$, is a birth-death process with parameters $\lambda_j = \lambda$ and $\mu_j = \mu\,(j \wedge s)$, $j \geq 0$. Assume the queue has traffic intensity $\rho = \lambda/\mu s \leq 1$, a necessary and sufficient condition for recurrence. Then from (5.3)

$$P_0\{M_1^+ > n\} = \left[\sum_{i=0}^{n} \pi_i^{-1}\right]^{-1}$$

$$= \left[\sum_{i=0}^{s} \left(\frac{\mu}{\lambda}\right)^i i! + \frac{s!}{s^s \rho^{s+1}} \sum_{i=0}^{n-(s+1)} \rho^{-i}\right]^{-1} .$$

Asymptotically, as $n \to \infty$

$$P_0\{M_1^+ > n\} \sim \begin{cases} (s^s/s!)n^{-1} , & \rho = 1 \\ (s^s(1-\rho)/s!)\rho^n , & \rho < 1 . \end{cases}$$

Thus for $\rho < 1$ we can use (2.16) to obtain

$$P_0\{\max_{1 \le j \le n} M_j^+ \le \ell\} \cong \exp(-e^{-a(\ell - b_n)}) \quad ,$$

where $a = \ln \rho^{-1}$ and $b_n = a^{-1} \ln(n\, s^s(1-\rho)/s!)$. Note that this is consistent with (4.3). ◀

6. Numerical Results

A simulation of the M/M/1 queue was performed to help assess the effectiveness of the basic procedures proposed in Section 3. Our goal is to estimate the d.f. of $\max_{1 \leq j \leq n} M_j^+$, where M_j^+ is either the maximum waiting time W, virtual waiting V, or number of customers Q in the system in cycle j. The processes $\{W_n : n \geq 0\}$ and $\{V_t : t \geq 0\}$ have continuous state spaces, $\{Q_t : t \geq 0\}$ has a discrete state space, and Λ is the appropriate extreme value distribution in all three cases. Using estimates a_n, b_n, we approximate $P\{\max_{1 \leq j \leq n} M_j^+ \leq x\}$ by $\Lambda((x + \delta - b_n)/a_n)$ where $\delta = 0$ for W_n and V_t, and we try $\delta = 0$ and $\delta = 1/2$ for Q_t. Theoretical values of a_n and b_n as well as the exact values of $P\{\max_{1 \leq j \leq n} M_j^+ \leq x\}$ are available from the results in Sections 4 and 5. All notation and conventions in this section are as in Section 3.

The random number generator used was the DEC-20 FORTRAN "RAN" function. In the terminology of Section 3, all experimental results reported are for basic linear or nonlinear regression procedures. Several values of k,n, and traffic intensity ρ were tested. Both the linear and nonlinear regression procedures were tried on W_n, while only the linear regression procedure was tried on V_t and Q_t. The nonlinear regression was performed using SUBROUTINE LSQFDN of the National Physical Laboratory Algorithms Library. Both full and partial delete as well as continuity corrections of $\delta = 0$ and $\delta = 1/2$ were tried on Q_t. For Q_t, a cutoff of $\epsilon = .0001$ was used. For W_n and V_t, a cutoff of $\epsilon = .001$ was used. Thus for W_n and V_t, i_0 is as follows:

k	n	i_0
100,000	2000	342
100,000	1000	686
50,000	1000	341
20,000	1000	135
50,000	250	801*
25,000	250	678
10,000	250	270

The starred entry indicates that for $k = 50{,}000$, $n = 250$, i_0 corresponds to $\epsilon = .01755$ rather than to $\epsilon = .001$. Note that in all cases, there were sufficient points for the regression for Q_t; and $y_{1,k},\dots,y_{i_0,k}$ were always distinct for W_n and V_t.

Tables 3-5 contain the results of the simulation for estimating a_n and b_n. Tables 6-10 contain the estimates of $P\{\max_{1\le j\le n} M_j^+ \le x\}$ by $\Lambda((x + \delta - b_n)/a_n)$, using the estimated values of a_n, b_n shown in Tables 3-5. The entries contained in the tables are the sample means of the various estimates over the number of replications and the half-length of a symmetric 90% confidence interval about the sample mean. For example, in Table 4, take $k = 50{,}000$, $n = 1000$, and 50 replications. Then a 90% confidence interval for a_{1000} based on 50 replications of the linear regression procedure is $[.1966 - .0039, .1966 + .0039]$. The corresponding true value of $a_{1000} = .2$. Tables 6-10 report the true values of $\Lambda((x - b_n)/a_n)$ and $P\{\max_{1\le j\le n} M_j^+ \le x\}$ for various x, as well as the corresponding values of $\Lambda((x + \delta - b_n)/a_n)$ using estimated a_n, b_n. For example, in Table 9, take $k = 50{,}000$, $n = 250$, and 100 replications. Then the true values of $\Lambda((9 - b_{250})/a_{250})$ and $P\{\max_{1\le j\le 250} M_j^+ \le 9\}$ are respectively $.7834$ and $.7831$, while the sample mean of $\Lambda((9.0 - b_n)/a_n)$ is $.7832$ with a corresponding 90% confidence interval half-length of $.0062$.

TABLE 3

Estimates of a_n and b_n for $\{W_n : n \geq 0\}$ in the M/M/1 queue with $\mu = 10$

k/n/#repl.	ρ	true values a_n	true values b_n	linear regression $\hat{a}_n$	linear regression $\hat{b}_n$	nonlinear regression $\hat{a}_n$	nonlinear regression $\hat{b}_n$
100,000/1000/50	.5	.2	1.1043	.1987 .0021	1.1069 .0035	.1981 .0028	1.1076 .0037
50,000/1000/50	.5	.2	1.1043	.1974 .0037	1.1043 .0055	.1969 .0050	1.1053 .0054
20,000/1000/50	.5	.2	1.1043	.1937 .0061	1.0993 .0094	.1913 .0079	1.1000 .0084
50,000/250/100	.5	.2	.8270	.1980 .0016	.8303 .0019	.1984 .0015	.8308 .0019
25,000/250/100	.5	.2	.8270	.1967 .0017	.8288 .0029	.1961 .0020	.8290 .0029
10,000/250/100	.5	.2	.8270	.1941 .0027	.8255 .0046	.1920 .0031	.8264 .0049
100,000/1000/50	.9	1.0	4.4998	.9717 .0115	4.4991 .0214	.9761 .0150	4.4980 .0214
50,000/1000/50	.9	1.0	4.4998	.9825 .0163	4.5324 .0276	.9840 .0228	4.5159 .0275
20,000/1000/50	.9	1.0	4.4998	.9748 .0252	4.5324 .0443	.9648 .0382	4.5245 .0460
100,000/2000/50	.9	1.0	5.1930	.9770 .0160	5.1784 .0292	.9714 .0232	5.1807 .0304

TABLE 4

Estimates of a_n and b_n for $\{V_t : t \geq 0\}$ in the M/M/1 queue with $\mu = 10$

		true values		linear regression	
k/n/#repl.	ρ	a_n	b_n	$\hat{a}_n$	$\hat{b}_n$
100,000/1000/50	.5	.2	1.2429	.1980 .0021	1.2451 .0034
50,000/1000/50	.5	.2	1.2429	.1966 .0039	1.2422 .0057
20,000/1000/50	.5	.2	1.2429	.1917 .0062	1.2371 .0093
50,000/250/100	.5	.2	.9657	.1977 .0016	.9692 .0018
25,000/250/100	.5	.2	.9657	.1966 .0018	.9673 .0029
10,000/250/100	.5	.2	.9657	.1946 .0027	.9648 .0047
100,000/1000/50	.9	1.0	4.6052	.9713 .0115	4.6050 .0214
50,000/1000/50	.9	1.0	4.6052	.9819 .0163	4.6308 .0276
20,000/1000/50	.9	1.0	4.6052	.9758 .0251	4.6401 .0442
100,000/2000/50	.9	1.0	5.0695	.9763 .0161	5.2840 .0293

TABLE 5

Estimates of a_n and b_n for $\{Q_t : t \geq 0\}$ in the M/M/1 queue with $\mu = 10$

k/n/#repl.	ρ	true values a_n	true values b_n	full delete linear regression $\hat{a}_n$	full delete linear regression $\hat{b}_n$	partial delete linear regression $\hat{a}_n$	partial delete linear regression $\hat{b}_n$
100,000/1000/50	.5	1.4427	8.9658	1.4350 .0409	8.9779 .0394	1.4358 .0405	8.9763 .0395
50,000/1000/50	.5	1.4427	8.9658	1.4224 .0519	8.9438 .0536	1.4224 .0530	8.9426 .0540
20,000/1000/50	.5	1.4427	8.9658	1.3783 .0611	8.8994 .0771	1.3801 .0609	8.8980 .0758
50,000/250/100	.5	1.4427	6.9658	1.4347 .0253	6.9853 .0179	1.4336 .0212	6.9842 .0180
25,000/250/100	.5	1.4427	6.9658	1.4224 .0288	6.9725 .0252	1.4179 .0301	6.9695 .0257
10,000/250/100	.5	1.4427	6.9658	1.4039 .0345	6.9491 .0379	1.3986 .0348	6.9437 .0374
100,000/1000/50	.9	9.4912	43.7087	9.2227 .1945	43.7424 .2232	9.2250 .1921	43.7421 .2226
50,000/1000/50	.9	9.4912	43.7087	9.3229 .2351	43.9337 .3107	9.3203 .2365	43.9316 .3116
20,000/1000/50	.9	9.4912	43.7087	9.1984 .3155	43.8827 .4561	9.2080 .3178	43.8893 .4580
100,000/2000/50	.9	9.4912	50.2875	9.2249 .2333	50.1689 .3108	9.2261 .2297	50.1684 .3085

TABLE 6

Estimates of $P(\max_{1 \le j \le n} M_j^+ \le x)$ for $\{W_n : n \ge 0\}$ in the M/M/1 queue with $\mu = 10$, $\rho = .5$

regression method	k/n/#repl.	$\Lambda((x-b_n)/a_n)/x$ $P(\max_{1 \le j \le n} M_j^+ \le x)$				
		.25/1.0390 .2493	.50/1.1776 .4996	.75/1.3535 .7499	.90/1.5544 .9000	.99/2.0243 .9900
linear	100,000/1000/50	.2457 .0052	.4971 .0072	.7489 .0062	.8997 .0037	.9900 .0007
nonlinear	100,000/1000/50	.2442 .0054	.4966 .0079	.7492 .0074	.8999 .0045	.9899 .0008
linear	50,000/1000/50	.2511 .0082	.5040 .0117	.7535 .0104	.9015 .0062	.9899 .0012
nonlinear	50,000/1000/50	.2489 .0077	.5035 .0121	.7539 .0116	.9014 .0074	.9896 .0015
linear	20,000/1000/50	.2626 .0136	.5193 .0194	.7638 .0174	.9053 .0107	.9897 .0022
nonlinear	20,000/1000/50	.2576 .0131	.5208 .0188	.7683 .0174	.9079 .0110	.9898 .0023
		$\Lambda((x-b_n)/a_n)/x$ $P(\max_{1 \le j \le n} M_j^+ \le x)$				
		.25/.7617 .2471	.50/.9003 .4986	.75/1.0762 .7496	.90/1.2771 .8999	.99/1.7471 .9900
linear	50,000/250/100	.2438 .0027	.4963 .0042	.7493 .0039	.9003 .0023	.9901 .0005
nonlinear	50,000/250/100	.2432 .0028	.4951 .0040	.7481 .0038	.8996 .0023	.9900 .0005
linear	25,000/250/100	.2464 .0042	.5002 .0060	.7525 .0052	.9020 .0030	.9903 .0005
nonlinear	25,000/250/100	.2454 .0045	.5001 .0060	.7531 .0053	.9025 .0032	.9904 .0005
linear	10,000/250/100	.2529 .0070	.5092 .0097	.7593 .0082	.9052 .0049	.9906 .0008
nonlinear	10,000/250/100	.2494 .0076	.5087 .0098	.7615 .0081	.9071 .0048	.9909 .0008

TABLE 7

Estimates of $P(\max_{1 \le j \le n} M_j^+ \le x)$ for $\{W_n : n \ge 0\}$ in the M/M/1 queue with $\mu = 10$, $\rho = .9$

regression method	k/n/#repl.					
		$\Lambda((x-b_n)/a_n)/x$ $P(\max_{1 \le j \le n} M_j^+ \le x)$				
		.25/4.1732 .2454	.50/4.8663 .4977	.75/5.7457 .7494	.90/6.7502 .8999	.99/9.1000 .9900
linear	100,000/1000/50	.2484 .0065	.5049 .0089	.7577 .0074	.9054 .0042	.9910 .0007
nonlinear	100,000/1000/50	.2492 .0065	.5050 .0089	.7569 .0079	.9045 .0049	.9907 .0008
linear	50,000/1000/50	.2415 .0079	.4952 .0116	.7490 .0101	.9001 .0059	.9900 .0010
nonlinear	50,000/1000/50	.2446 .0082	.4990 .0117	.7514 .0111	.9006 .0069	.9897 .0012
linear	20,000/1000/50	.2422 .0131	.4968 .0184	.7491 .0161	.8991 .0094	.9894 .0017
nonlinear	20,000/1000/50	.2442 .0132	.5045 .0201	.7558 .0191	.9012 .0119	.9888 .0023
		$\Lambda((x-b_n)/a_n)$ $P(\max_{1 \le j \le n} M_j^+ \le x)$				
		.25/4.8663 .2477	.50/5.5595 .4989	.75/6.4389 .7497	.90/7.4433 .9000	.99/9.7931 .9900
linear	100,000/2000/50	.2552 .0090	.5100 .0121	.7590 .0101	.9049 .0059	.9906 .0010
nonlinear	100,000/2000/50	.2539 .0092	.5112 .0132	.7607 .0119	.9055 .0070	.9905 .0012

TABLE 8

Estimates of $P(\max_{1 \le j \le n} M_j^+ \le x)$ for $\{V_t : t \ge 0\}$ in the M/M/1 queue with $\mu = 10$,

using linear regression

k/n/#repl.	ρ	$\Lambda((x-b_n)/a_n)/x$ $P(\max_{1\le j\le n} M_j^+ \le x)$				
		.25/1.1776 .2493	.50/1.3162 .4996	.75/1.4921 .7499	.90/1.6930 .9000	.99/2.1630 .9900
100,000/1000/50	.5	.2459 .0050	.4982 .0072	.7503 .0062	.9006 .0037	.9901 .0007
50,000/1000/50	.5	.2523 .0082	.5062 .0122	.7555 .0111	.9025 .0067	.9900 .0012
20,000/1000/50	.5	.2631 .0134	.5224 .0196	.7676 .0174	.9077 .0106	.9902 .0020
		$\Lambda((x-b_n)/a_n)/x$ $P(\max_{1\le j\le n} M_j^+ \le x)$				
		.25/.9003 .2471	.50/1.0390 .4986	.75/1.2148 .7496	.90/1.4157 .8999	.99/1.8857 .9900
50,000/250/100	.5	.2430 .0027	.4959 .0040	.7494 .0039	.9005 .0025	.9901 .0005
25,000/250/100	.5	.2466 .0042	.5005 .0062	.7528 .0054	.9021 .0032	.9904 .0005
10,000/250/100	.5	.2518 .0070	.5076 .0097	.7579 .0082	.9044 .0049	.9905 .0008
		$\Lambda((x-b_n)/a_n)/x$ $P(\max_{1\le j\le n} M_j^+ \le x)$				
		.25/4.2785 .2454	.50/4.9717 .4977	.75/5.8511 .7494	.90/6.8555 .8999	.99/9.2053 .9900
100,000/1000/50	.9	.2481 .0065	.5048 .0089	.7576 .0074	.9054 .0042	.9910 .0007
50,000/1000/50	.9	.2414 .0079	.4953 .0116	.7491 .0100	.9002 .0059	.9900 .0010
20,000/1000/50	.9	.2414 .0129	.4958 .0184	.7484 .0159	.8987 .0094	.9893 .0017
		$\Lambda((x-b_n)/a_n)/x$ $P(\max_{1\le j\le n} M_j^+ \le x)$				
		.25/4.9717 .2477	.50/5.6648 .4989	.75/6.5442 .7497	.90/7.5487 .9000	.99/9.8985 .9900
100,000/2000/50	.9	.2550 .0090	.5100 .0117	.7591 .0102	.9050 .0059	.9906 .0010

TABLE 9

Estimates of $P(\max_{1 \le j \le n} M_j^+ \le x)$ for $\{Q_t : t \ge 0\}$ in the M/M/1 queue with $\mu = 10$, $\rho = .5$, using full delete linear regression, with continuity correction $\delta = 0$

k/n/#repl.	$\Lambda((x-b_n)/a_n)/x$ $P(\max_{1 \le j \le n} M_j^+ \le x)$				
	.3766/9 .3761	.6137/10 .6135	.7834/11 .7833	.9408/13 .9408	.9924/16 .9924
100,000/1000/50	.3769 .0106	.6153 .0129	.7839 .0114	.9394 .0057	.9915 .0013
50,000/1000/50	.3890 .0144	.6264 .0173	.7908 .0151	.9407 .0077	.9913 .0020
20,000/1000/50	.4073 .0213	.6447 .0240	.8032 .0198	.9443 .0092	.9916 .0022
	$\Lambda((x-b_n)/a_n)/x$ $P(\max_{1 \le j \le n} M_j^+ \le x)$				
	.3766/7 .3744	.6137/8 .6128	.7834/9 .7831	.9408/11 .9408	.9924/14 .9924
50,000/250/100	.3732 .0048	.6130 .0065	.7832 .0062	.9399 .0032	.9918 .0008
25,000/250/100	.3784 .0066	.6186 .0088	.7872 .0080	.9410 .0040	.9918 .0010
10,000/250/100	.3870 .0101	.6273 .0118	.7933 .0101	.9427 .0050	.9919 .0013

TABLE 10

Estimates of $P(\max_{1 \le j \le n} M_j^+ \le x)$ for $\{Q_t : t \ge 0\}$ in the M/M/1 queue with $\mu = 10$, $\rho = .9$, using full delete linear regression, with continuity correction $\delta = 0$

k/n/#repl.	$\Lambda((x-b_n)/a_n)/x$ $P(\max_{1 \le j \le n} M_j^+ \le x)$				
	.2644/41 .2599	.5293/48 .5272	.7604/56 .7599	.9089/66 .9088	.9906/88 .9906
100,000/1000/50	.2620 .0070	.5346 .0106	.7678 .0097	.9133 .0059	.9911 .0012
50,000/1000/50	.2578 .0095	.5275 .0144	.7604 .0129	.9084 .0074	.9901 .0013
20,000/1000/50	.2628 .0147	.5345 .0209	.7647 .0179	.9095 .0104	.9898 .0020

	$\Lambda((x-b_n)/a_n)/x$ $P(\max_{1 \le j \le n} M_j^+ \le x)$				
	.2801/48 .2779	.5085/54 .5074	.7695/63 .7693	.9035/72 .9034	.9901/94 .9900
100,000/2000/50	.2861 .0104	.5204 .0142	.7796 .0124	.9085 .0077	.9903 .0015

Here are some general observations on Tables 3-10. The approximation of $P\{\max_{1\leq j\leq n} M_j^+ \leq x\}$ by the exact value of $\Lambda((x - b_n)/a_n)$ improves with greater x, greater n, and lesser ρ, and is about equally good for W_n, V_t, and Q_t. Note that confidence intervals generally cover the true values (excepting the use of continuity correction $\delta = 1/2$ for Q_t) especially in the percentile evaluations of Tables 6-10. The bias and confidence interval widths tend to decrease as k increases, and decrease as x increases in the percentile evaluations. The results for $\rho = .9$ are just about as good as for $\rho = .5$. The results for Q_t are almost as good as for W_n and V_t, except that the confidence interval widths are greater for Q_t, which can be largely attributed to the greatly reduced number of regression data points.

For the W_n process, the linear regression procedure tends to produce estimates with slightly smaller biases and confidence intervals than does the nonlinear regression procedure. Although neither is a clear winner, in this case the linear regression would appear to be better, if just for the reason of the greater ease and lesser expense it entails. (However, the nonlinear regression problem under consideration is a nice one, and the linear regression procedure could certainly provide good starting values of a_n, b_n for the nonlinear regression.)

For the Q_t process, the (null) continuity correction of $\delta = 0$ is superior to $\delta = .5$; and thus results for $\delta = .5$ are not shown. This is no surprise in the current case since the approximation of the exact d.f. of $\max_{1\leq j\leq n} M_j^+$ by the value of Λ using exact a_n, b_n has the same error patterns for Q_t as for W_n and V_t. The full and partial delete procedures give comparable results, so only the estimates of a_n, b_n are shown for the partial delete procedure.

Some experimentation on the W_n process (not reported in the tables) using linear regression estimates a_n, b_n to arrive at estimates $a_{n'} = a_n$, $b_{n'} = a_n \ln(n'/n) + b_n$ and corresponding percentile evaluations for n' cycles was performed using $n' = 1000$ and $n = 250$ and 100. The general pattern of results is that for fixed k, smaller values of n give smaller confidence intervals but greater biases, unless k is very small. The decrease in confidence interval widths is mostly attributable to the use of more regression data points for smaller n. Again considering the results reported in the tables for $n = 1000$, in many instances $k = 50{,}000$ performs better than $k = 100{,}000$. All this suggests that increasing the k/n ratio beyond a certain point may not be worthwhile. However, optimal values of k and n are very problem dependent.

Most of the cost in using these procedures derives from the development and running of the simulation model. Another substantial portion of the cost is keeping track of the $y_{i,k}$'s. However this can be done once, and then empirical distribution functions corresponding to the different values of n can easily be constructed. Different variations of the regression procedures could be performed for each value of n. In short, the variations are not competitors, but rather can be used to help assess the validity of the results. Any special insight to the problem at hand could also be used to advantage.

Acknowledgement

This research was supported by National Science Foundation Grants MCS75-23607 and MCS79-09139 and Office of Naval Research Contract N00014-76-C-0578. We gratefully acknowledge the assistance of P. Glynn, P. Heidelberger, and L. Jow in carrying out numerical computations for an earlier version of this paper. Finally, our thanks goes to several referees who suggested numerous improvements to the manuscript.

REFERENCES

[1] ANDERSON, C.W. (1970). Extreme value theory for a class of discrete distributions with applications to some stochastic processes. J. Appl. Probability 7, 99-113.

[2] BERMAN, S. (1962). Limiting distribution of the maximum term in sequences of dependent random variables. Ann. Math. Statist. 33, 894-908.

[3] CHUNG, K.L. (1960). Markov Chains with Stationary Transition Probabilities. Springer-Verlag, Berlin.

[4] COHEN, J.W. (1969). The Single Server Queue. North Holland, Amsterdam.

[5] EPSTEIN, B. (1960). Tests for the validity of the assumption that the underlying distribution of life is exponential, Part I. Technometrics 2, 83-101.

[6] FISHER, R.A. and TIPPETT, L.H.C. (1928). Limiting form of the frequency distribution of the largest or smallest member of a sample. Proc. Cambridge Phil. Soc. XXIV, Part II, 180-190.

[7] GNEDENKO, B.V. (1943). Sur la distribution limite du terme maximum d'une serie aléatoire. Ann. Math. 44, 423-453.

[8] de HAAN, L. (1970). On Regular Variation and Its Application to the Weak Convergence of Sample Extremes. Mathematisch Centrum, Amsterdam.

[9] HEYDE, C.C. (1971). On the growth of the maximum queue length in a stable queue. Operations Res. **19**, 447-452.

[10] IGLEHART, D.L. (1972). Extreme values in the GI/G/1 queue. Ann. Math. Statist. **43**, 627-635.

[11] IGLEHART, D.L. (1977). The regenerative method for simulation analysis. *Current Trends in Programming Methodology*, **3**: Software Modelling and Its Impact on Performance. M. Chandy and R. Yeh, eds., 52-71, Prentice-Hall, Englewood Cliffs, N.J.

[12] MANN, N.R., SCHAFER, R.E., and SINGPURWALLA, N.D. (1974). Methods for Statistical Analysis of Reliability and Life Data. John Wiley and Sons, New York.

[13] National Physical Laboratory Algorithms Library Reference No. E4/18/F, Teddington, Middlesex, England.

[14] PICKANDS, J. III (1975). Statistical inference using extreme order statistics. Ann. Statist. **3**, 119-131.

[15] WEISSMAN, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. J. Amer. Statist. Assoc. **73**, 812-815.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 64	2. GOVT ACCESSION NO. AD-A119119	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Regenerative Simulation for Estimating Extreme Values		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) DONALD L. IGLEHART MARK L. STONE		8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0578
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Operations Research Stanford University Stanford, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (NR 042-343)
11. CONTROLLING OFFICE NAME AND ADDRESS Statistics and Probability Program Office of Naval Research (Code 436) Arlington, Virginia 20360		12. REPORT DATE JULY 1982
		13. NUMBER OF PAGES 42
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release: distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

An earlier version of this paper was issued as Technical Report No. 43 in October 1977.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

EXTREME VALUES, REGENERATIVE METHOD, SIMULATION, STATISTICAL ESTIMATION

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

SEE ATTACHED

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

"Regenerative Simulation for Estimating Extreme Values", Technical Report No. 64

Let $\{X_t : t \geq 0\}$ denote the regenerative process being simulated and assume that X_t converges weakly (in distribution) to a limit random variable X. Our conern in this paper is in estimating the extreme values of the process $\{X_t : t \geq 0\}$. Suppose we are interested in the largest value attained in the interval $[0,t]$: $X_t^* = \sup\{X_s : 0 \leq s \leq t\}$. Examples of this are the maximum queue lengths or waiting times in a queueing system. As t increaees so will X_t^* without bound if the state space of $\{X_t : t \geq 0\}$ is unbounded. This report develops several methods for estimating the distribution of X_t^*. When the regenerative process is either the GI/G/1 queue or a birth-death process theoretical results are available for the distribution of X_t^*. The waiting time, queue length, and virtual waiting time for an M/M/1 queue were simulated. The methods for estimating the distribution of X_t^* were employed and the simulation results compared with the theoretical results.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)